R E P O R T R E S U M E S

ED 003 908 48

SOME RECOMMENDATIONS ON THE SELECTION OF ENTRIES FOR A BILINGUAL DICTIONARY.

BY- SWANSON, DONALD C.

REPORT NUMBER NDEA-VI-16-7 PUB DATE 60

EDRS PRICE MF-$0.09 HC-$0.64 16P.

DESCRIPTORS- *MODERN LANGUAGES, *DICTIONARIES, *BILINGUALISM, *ENGLISH, BLOOMINGTON, INDIANA

A GUIDE FOR PREPARING BILINGUAL DICTIONARIES WAS DEVELOPED, EMPHASIZING PROPER METHODS FOR SELECTING DICTIONARY ENTRIES. ASSUMED PURPOSES OF BILINGUAL DICTIONARIES WERE CONSIDERED IN CHOOSING THE SELECTION CRITERIA. A LIST OF POSSIBLE CRITERIA WAS PRESENTED, CENTERING ON THE FORM, SYNTAX, AND MEANING OF WORDS CONSIDERED FOR SELECTION AND THE FREQUENCY OF WHICH THESE WORDS ARE USED. IT WAS SUGGESTED THAT BILINGUAL DICTIONARIES OF INDO-EUROPEAN LANGUAGES SHOULD CONTAIN A MINIMUM OF 5,000-10,000 ENTRIES. THIS REPORT IS ONE OF A SERIES OF 13 PAPERS PRESENTED AT A CONFERENCE ON LEXICOGRAPHY, INDIANA UNIVERSITY, NOVEMBER 11-12, 1960. (JH)

ED003908

16-7

SOME RECOMMENDATIONS ON THE SELECTION OF ENTRIES FOR A BILINGUAL DICTIONARY

Donald C. Swanson
U. of Minnesota

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFA
Office of Education
This document has been reproduced exactly as received fro
person or organization originating it. Points of view or op
stated do not necessarily represent official Office of Edu
position or policy.

A. Assumptions

A bilingual dictionary can be useful and desirable to several kinds of people; students, travelers, translators, and linguists. The last group is, and will be in the foreseeable future, the smallest, and we may as well face the fact that only a small percentage of potential users of bilingual dictionaries (at least of those of major languages) will be linguistically alert.

There can be two principal purposes of a bilingual dictionary, the customary one of reference for the student (of either language); the other that of a ready reference guide for the linguist who is not a specialist, and to whom each half of the dictionary serves as an index for the other.

In this paper, remarks will be based chiefly on the Indo-European (IE) languages. A bilingual dictionary composed between almost any two of these will be easy to set up because of similar grammatical and syntactic categories and word-classes, but of course caution is needed even here, since translation or equivalence is not a clue to classification. A bilingual dictionary can retain for the general public, much of the traditional approach, but the more firmly established techniques of recent linguistics can be unobtrusively adopted.

The number of entries (words; see below) necessary for usefulness is anybody's guess. I suggest for English and many other IE languages 5000 as a bare minimum and perhaps 10,000 as a working solution. Above this figure there is a drop-off in utility, and the extra entries tend to pad the work with more word-lists. A presentation of word-formational principles (see below, B-3) in IE and certain other language families can serve to account for many absent entries.

According to my records on bilingual dictionaries of IE dialects, English and German have developed the most productive traditions in dictionary making, followed at a distance by French, Russian and Italian, in that order[1]. Since the number of English speakers (and of those to whom English is a well-understood second language) is very large, the established tradition is likely to continue. Because of recent improvements in our understanding of English structure and the possibility of still further analysis of English for the purpose of squeezing out all its translation powers we can now make it as useful as it can be as a translating and indexing language, in other words not just a tool for English speakers but a kind of international language for general lexicography. English is not (as of course no natural language is) suited for facile interpretation of all kinds of structures; it is loosely Germanic in structure and partly bound to one environment. But, by exhausting all its native possibilities, then utilizing the tradition of 'translation English' (long in use to handle Greek and Latin), and finally the scientific vocabulary and the international loanwords, we may arrive at a practicable ideal. This international English will have to be defined as a combination of American and British, the main varieties[2], for in practice Americans have used American words and usage, and the English have used British, each without a thought for the other.

B. Selection of Entries

Entries will be, as conventionally, primarily words. Current linguistic analysis provides us with more precise (though not yet definitive) criteria for "words". For IE languages in general there is no important problem either in the nature of word or of its categories. For a language like English, phonetic criteria[3] may determine the words (usually but not always coinciding with spelled words); for 'archaic'[4] IE languages

morphology will have to suffice. A typical archaic feature is the group of 4 enclitics (a variety of quasi-word) in Latin, namely -que, -ve, -dum, -ne. They have no stress and may not stand alone; but they are not suffixes and they may in part be substituted by free forms (et, --, vel, nonne or num). They will not satisfy Francis' definition[5](based undoubtedly on English): "...a linguistic form consisting of a base and a superfix.." Hockett's definition (A Course.., N.Y., 1958, p. 167), although also not considering this particular problem, could conceivably embrace the group: "...any segment ... bounded by a succession of points at which pausing is possible." (We may invent a distinction between full pauses and half pauses.)

Nida[6] reminds us of words based on phrases whose second constituent has become atonic; kind-of, sort-of, used-to illustrate the type, and these words should be entered if only to call attention to their real existence.

Another type of main entry to be considered (on the English side of our dictionary) consists of a group of nominal compounds or phrases (the phrase being composed of adjective + noun or adjunct noun + noun): adhesive tape hiding place, carbon paper. Such words (or phrases) need not be sub-entered under their constituents. A word like no one is often missing as a lexical main entry (and even as a subentry) because the orthographic space gives it the appearance of a phrase.

Assuming the impossibility of a complete bilingual dictionary, I give a list of criteria to use in realistic selection of entries, based on the fact that words have form, syntax, meaning and a relative frequency.

1. relative frequency. We may judge the most frequent words (and morphemes) the most complicated in form or meaning and therefore the most needy of elaborate treatment in a bilingual dictionary. Only a few languages have been provided with frequency lists, namely American English, Classical Latin, Greek (Homeric, New Testament), Arabic, German, French, Spanish, Brazilian Portuguese, Swedish[7]. We need such tallies for all

languages. There also exist indices verborum for a variety of texts in some other languages but these have only substitute value. Several of the older frequency lists were of word-forms (that is inflected words); in addition certain naive practices have spoiled the usefulness of some lists[8], such as failure to distinguish sandhi variants. In the Thorndike-Lorge list of 1000 most frequent "words" occur 63 forms which can be deleted as grammatical variants: 4 plurals (feet etc), 3 gerunds, 2 comparatives, 1 sandhi variant (a), 5 phrases (don't etc), and 48 preterites (including fell; but fall is not in the list). In their original 1000 (given on pp. 270-274) occur 104 words[9] not found in the later list. One may (pending the compilation of a scientific list) use these cast-offs to replace the 63 false inclusions; this gives us a net total of 1041 most frequent American English words based on a count of written documents.

2. "function-words". The particles (traditional prepositions, conjunctions, and adverbs) which constitute the syntactic cement of discourse must be completely entered and given the necessary treatment, even to the extent of making this lexical group an index of particle syntax. The two recent works of Nida and of Francis[10], despite their different approaches to the subject, give us handy classifications and lists of this lexical area for English. It may be noted that Nida's treatment of adverbs is fuller than Francis'.

3. word-formation. In a bilingual dictionary there should be a fairly detailed essay (either in the preface or as an appendix) on word-formational habits (and presumably most languages have devices for forming derived words from morphemes or prime words). This essay is also ideally present in a grammar of a language but in practice many grammars (even recent ones) ignore or minimize word-formation[11]. Such a guide will cut down the need for extensive lists whose synchronic etymology (derivability) is obvious

and simple, as e.g. English goodness, wholesomeness and the like. Some items in a derivational paradigm may be included for other reasons, of course. A preliminary study of derivation will have yielded the processes by which 'new' words are made. Generalizing only for the IE type, I may name the following four major categories:

a. compounding (composition) occurs in either prefixal or 'co-valent'[12] types and will be found in most word-classes.

b. suffixation: found in most word-classes but showing extreme variation in the noun and adjective.

c. zero-change (Bloch-Trager, 59 & 63-64): new formation by shift of a word from one class to another, with or without other changes including ablaut, but not including formative suffixation or compounding. Denominative verbs and deverbative nouns are typical cases.

d. reduplication: rare in archaic IE, mildly productive in some recent IE languages.

One will be on guard in equating English with foreign word-formational devices. For example Eng. -ish (in roddish) comes out as a prefix in Latin sub- (in subrufus). Formative affixes of these types are easily included as main entries.

It appears that in general prime words have a higher frequency in actual discourse than derived words. For instance, in the Thorndike-Lorge list, of the first 500 words only 19 (3.8%) are derived (e.g. any/thing); but in the second 500 words 35 (or 7%) are derived from primes. Obviously as the relative frequency goes down the percentage of derived words rises[13].

In our foreign language we must look for various kinds of formatives, but we must not be surprised to find gaps of one kind or another. For example, although a vast majority of IE languages show 'diminutive' suffixes, there appear to be none in Modern Indic[14]; and compounds (co-valent type)

occur almost universally in IE, but Pashto seems to show few or none[15]. Relative frequency of type is also unpredictable: modern English has but few dvandvas (outside of scientific terms) but modern Greek and Hindustani have many examples of this compound type.

While word-formational patterns will give us rules and obviate the necessity of including vast numbers of rare or derived words, we must watch for those which must be included because of morphophonemic peculiarities[16] or unexpected meanings (e.g. shifted or metaphoric: German Hochzeit, Lat. comprehendo, Eng. buttercup, heavy-weight).

I may summarize at this point a study intended to show what the examination of word-formation in one text of a language may reveal when one has in mind the lexical structure of that language[17]. Lucretius, the Epicurean poet of the first century B.C., has a total of 4295 different words, excluding sandhi and grammatical variants. Of the total, 1457 are prime words, and 2838 are derived words (ratio 1 : 2); in his text of 7415 lines of dactylic verse (about 240 printed pages) two of every three different words are derivable from primes by compounding, suffixation or zero-change. The breakdown (by processes) of the derived words in Lucretius is: suffixed 1065 items, compounded 1088, zero-change 683; in addition there are two cases only of reduplication (quisquis; quantus...quantus, a phrase). Since the number of root-words (primes) in any IE language is finite, we may tentatively conclude that for any text of a language, the greater the bulk the lower the percentage of primes to derived forms.

4. semantic criteria. A passage in Bloomfield's Language (p. 140: "...the weak point in language-study...") has caused undue pessimism among many students of language. In daily practice we have to deal with meanings, above all in a dictionary[18]. I suggest two alternative techniques for handling many words and morphemes. Assume that we have first isolated and

classified words on a morphological basis, we can then attempt a:

a. classification by contrasts("What is the antonymous expression of ...?")

light rain (snow, dumbbell) : heavy rain (snow, dumbbell)

light complexion (color, red) : dark complexion (color, red)
deep blue (etc)

or

wide door (board : narrow door (board)

wide array : small array

wide difference (of opinion) : slight difference

or

dry towel (street) : wet towel (stree')

dry wine (martini) : sweet wine (--)

dry skin : oily skin

b. classification by substitution: If we are confronted by an isolated utterance, "He walked up to the bar", we can determine from this sentence (the 'inner context') only that bar is a noun. If we consult the 'outer context' (the paragraph or the situation), we determine that bar is, by substitution, either a 'counter for serving drinks' , or 'a barbell', or a 'sandbar', etc. The number of possible contrasts and substitutions gives us a rough notion of the number of meanings in the same form. We may, of course, choose to consider some of the possible meanings as impractical for our dictionary on the grounds that they are vestigially archaic, slangy or otherwise of little use.

Homonyms should usually be distinguished, especially if they show differences in morphology or syntax (congruence or word-order). Thus Lat. multa 'many things' and multa 'fine (N) coincide phonologically only at one point.

Although the procedure may be difficult to justify, we should set up

semantic categories as a guide to selecting useful entries (useful despite the relative infrequency of some members of such categories). Compilers such as Roget, Soule, and C.D. Buck have given us convenient lists to work from. But all three have started from English only or primarily and some adjustments would be necessary to allow a more universal approach[19]. A potentially helpful guide in this instance, although prepared for cultural anthropologists, is the Outline of Cultural Materials, prepared by G. Murdock (New Haven, 1945[2]).

In the special area of plants and animals a general bilingual dictionary cannot attempt to give the complete rundown one would expect in a technical botanical/zoological lexicon. The well-known varieties and local items (farmers and rural dwellers will have a greater knowledge than city-dwellers) can probably be limited to a total of 100 or 200. Folk-terms and scientific terms for plants and animals are often at variance (sometimes by choice, as a result of language migration). Thus the American robin (actuallya thrush) resembled the English robin to early settlers, and our buffalo (really bison) resembled an animal of the eastern Hemisphere.

Favorite expressions may sometimes be cast in patterns and make us seek other, comparable patterns. An English pattern which comes to mind is the substitution in a half-dozen or more phrases or compounds of a metaphor for the qualifier very. Thus, instead of very high we may prefer sky-high and, for very deaf, stone-deaf[20]. A too mechanical approach to lexicography may cause us to overlook legitimate data, and this is one area where we can encourage our informants (of the foreign language) to produce semantic paradigms when we suspect that they exist. And in like fashion it does not seem to be a dangerous practice to inquire about metaphorical meanings[21] in those words whose primary meanings are animal names, body parts and color terms.

5. feed-back from the foreign-to-English part. Assuming that we begin work on a bilingual dictionary by eliciting foreign equivalents of English sentences, phrases and words (in a variety of simple contexts), we will occasionally encounter foreign items not provided for in the English part. Thus, if we gt only Span. remoto as an equivalent to Eng. remote we may get, in another context, Span. rincón 'inner angle', 'outer angle', but also 'remote locality','out-of-the-way place'. The English entries hence become an index to unexpected foreign items or ones not easily discoverable.

By paying continuous attention to morphemic analysis we may get another kind of feed-back. Many German-English dictionaries translate Wirt by 'landlord; owner; host'. If we notice in other contexts (or from our own lexical data), such compounds as Gastwirt, Landwirt, we are justified (although this morpheme is rare finally) in adding the meaning '--keeper, --tiller'. In this connection it is important to have at hand a reverse-index of the foreign langu ge[22] to check the relative distribution of final compositional or suffixal morphemes. With English instead of (archaic in (my) stead) may be associated, by some cross reference, a final element -stead in bedstead, roadstead, homestead.

6. cultural items. A culture (defined here as a more-or-less unified linguistic area) is not aware that its facets are different, peculiar, or even characteristic, unless so informed by an outsider or unless confronted suddenly by another culture. A part of the business of a bilingual dictionary is to recognize lexic l items reflecting these differences nd to treat them accordingly, that is not to minimize them by seeking an exact (often artificial) equivalent in the other language. Thus king will translate ancient Greek basileús only if accompanied by a cultural note explaining the differences in connotation.

The dictionary must be careful to include older items in a continuing

culture even where, as in many parts of the USA, technology has replaced many traditional things and concepts. The old may not only be encountered in literature (history, novels, ballads) but it may yet live on residually in the most advance societies or be present among out of the way members of the culture. The most obvious lexical area here is the realm of handicrafts and home manufacture, such as weaving, pottery, working with leather, metals, wood. We can neglect neither the "latest terms" of science nor the older terms for an older technology. But there are other lexical items, more or less unpredictable, and the following lists indicate the better known examples, and ones which will be suggestive of still other items. The terms (and the things or concepts) may not be structurally significant[23] in the culture, but they may be contrastively interesting and important as regards their treatment in a bilingual dictionary.

American English: attic, blizzard, ice-cream, cocktail, bath-room, date (V), the West, coffee-break, drug-store, boy friend, hitch-hike, week-end.

Latin: ars, imperium, pius, auctoritas, virtus, officium, religio, fas, cliens, rex, patria.

Ancient Greek: hybris, atē, logos kosmos, nemesis, daimon, polis, paideia, aretē, gnosis, sophrosynē, xenos.

Modern Greek: onomasía, papa'dja, 'glendi, fi'lotimo, 'kalanda (pl), kafe'nio, kombo'loi, ta'verna.

Sanskrit: dhárma, ártha, kárman, kāma, brahmán, ṛṣi.

Old Italian: gentile, donna, cotera, innamorato, corteggieno, piazza, uomo-di-corte

0. A Proposal

Everyone who has worked with informants has grown impatient with the time-waste involved in writing English equivqlents. An official or semi-official alphabetized check-list of English entries and sub-entries could be easily prepared by an ad hoc committee of the LSA, perhaps in conjunction

with a British committee. A preliminary compilation could then be mailed out to volunteer members of the LSA, the MLA and the AAA for criticism and addenda; by this means accidental oversights and regional variants[24] could be gathered in by the committee for consideration. The final publication, mimeographed or multilithed, would provide a definitive workbook for all lexical purposes. If each page is doubled, the carbon copy would easily be used for reversing the entries. An appendix of semantic categories[25] would relieve the strain on the memory of both the interviewer and the informant.

Notes

1. Of 84 extinct and extant IE languages there are 62 English bilingual dictionaries, good, bad or indifferent; for the same group there are 46 German dictionaries, 23 French, 15 Russian and 10 Italian. (The records are probably incomplete but also probably representative of the real ratios.) Among the 22 languages for which there are no English dictionaries the minor Germanic, Slavic and Romance dialects constitute a majority. Many of the English-to-IE dialect dictionaries are poor or out-of-date and need redoing. Furthermore, there is no reason why regional dictionaries could not be made of e.g. English-Andalusian, English-Sardinian, English-Pomoranian, English-Frisian.

2. A master-list of English entries could accomodate varieties of both regions by simple abbreviation marks (A, B). This is standard practice in better German dictionaries where standard and S (Austria, etc) are usually distinguished, and in Spanish dictionaries where Iberian and American varieties are listed. The researches of A.W. Read would be extremely valuable for the British part of our entries.

3. Including the (sometimes sole) criterion of open juncture which is the same as unfilled phonetic space.

4. "Archaic" here refers to the type of grammatical (esp. inflectional) format associated with Greek, Latin and Sanskrit, and usually labelled synthetic. "Recent" refers to the other end of the (historical-typological) pole and includes a variety of disparate items usually labelled analytic. Archaic-type languages are mostly extinct, but archaic features remain in some extant languages in varying proportions.

5. W.N. Francis, "The Structure of American English" (N.Y., 1958), p. 204.

6. E. Nida, "Synopsis of English Syntax" (Norman, 1960), p. 138.

7. The Swedish list, by M. Allwood & I. Wilhelmsen (Rock Island, 1947), is not based on a count of Swedish texts but is a translation of Eaton's summary! The fact is that few of the frequency lists are satisfactory. Attention should be called, finally, to M. West's "General Service List of English Words" (London, 1953), dealing with semantic frequencies in British usage.

8. For example, in Paul B. Diederich's Chicago diss., "Frequency of Latin Words.." (1939), short-cuts (or misapprehensions) have caused the following pairs of words to be counted together: qui and quis, hic (pron.) and hic (adv.), cum (prep.) and cum (conj.), but nec and neque (sandhi variants) are counted separately.

9. For the record these additional 104 items are: (Thorndike, E.L. & Lorge, I.-"Teacher's Word Book of 30,000 Words", N.Y., 1944)

apple	cake	fancy	jump	proper	shut
attend	cap	fat	lesson	proud	silk
asleep	careful	feed	lion	pure	slow
band	choose	fence	list	quick	sometime
basket	Christmas	firm	load	quiet	strike
beast	clock	fix	loud	rapid	tail
bee	cloth	flow	mail	roof	teacher
bell	clothe (-ing)	former	mill	row	thick
bless	coast	fourth	moon	sad	throw
blind	corn	gentle	narrow	sand	tongue
bone	cow	gift	nest	seed	unite
bottom	crown	grace	noise	seize	wash
bow	delight	grain	noon	(self)	waste
brave	divide	grave	oak	separate	weak
breakfast	dust	hide	path	shake	wheat
broad	entire	hunt	pen	sheep	wheel
brook	equal	journey	perfect	shine	yesterday
butter	extend				

10. Nida (op.cit.) lists and discusses limiting substitutes (p. 52), clause markers (p. 54), coordinators (p. 59), antecedents (p. 93f), conjunctive markers (p. 190f). W.N. Francis (op.cit.) gives prepositions (and prepositional phrases) on pp. 306-308 (a total of 76), 16 adjective qualifiers (p. 278; to which add the colloquial real), coordinators (and, or) on p.279, includers (of included clauses) on p. 390 (about 46 in number, e.g. after, as...as, till).

11. Francis is interested primarily in analysis; one must consult his scattered paragraphs (e.g. pp. 208, 240, 260) for word-formation. Except for isolated articles on specific affixes and other processes in English word-formation, and the 6th volume of Jespersen's antiquated "Modern English Grammar", the subject as a whole is awaiting systematic treatment. Of 4 US Army dictionaries produced (as TM's; now published) at the end of the war, the Russian-English and the Spanish-English, despite otherwise elaborate grammatical treatments of their specific languages, omit word-formation; the German one devotes about 7 pages; and the Chinese nothing, though one would not expect much. Word-formation is the one subject I expect to see in both grammar and dictionary.

12. "Prefixal" is self-explanatory; "co-valent" is used (instead of non-prefixal) to refer to all other types collectively, since many of them contain noun + noun, adjective + adjective, etc. The classification is based primarily on archaic IE languages. (The two types seem to be equal in bulk.)

13. In Diederich's list (op.cit. in note 8) the first hundred Latin words have 15 derived, the second hundred 17, the third hundred 37 derived words.

14. I have been unable to elicit any from speakers of Sindhi, Hindi, Oriya or Maithili.

15. According to H. Penzl's "A Grammar of Pashto" (Wash., D.C., 1955) p. 68. The lack of compounds in Pashto is a major deviation. Kurdish, on the other hand (E.N. McCarus, "Kurdish Grammar", ibid., 1958, 83-87) has the "normal" types of compounds.

16. Formations like baker, carver, starter and numerous other agents/instruments in -er need not be fully listed, but morphophonemic exceptions like lawyer (and sawyer ?) should be entered. Important suffixes can be entered with suitable statistical and semantic information. For example, the English suffix -age (about 75 current formations in American English) shows three main semantic groups:

a. quantity/bulk: acreage, dosage, silage
b. residence: hermitage, orphanage, parsonage
c. status (rare): marriage, parentage

17. Lexical structure can be determined only in part from a dictionary, but it can be determined fairly precisely from a sampling of texts, and the resulting information is useful in compiling a dictionary. For such a sampling of ancient Greek see Dorothy M. Clay's University of Minnesota dissertation "A Formal Analysis of the Vocabularies of Aeschylus, Sophocles and Euripides", of which Part II (Classified Lists) was published in 1958, and Part I (Intro. and Analysis) is imminent. This is the answer to D.S. Worth's query (Word 16, 1960, 277f): "...the basic and still unsolved problem is to discover the systematic framework in terms of which the lexicon can be described." The ratio of derived to underived words in Greek tragedy is about 4 : 1; Euripides, whose bulk is about 3 times that of Sophocles, shows, of the total, 77% derivatives, Sophocles shows 75%.

18. Bloomfield (p. 140) lists 4 types of definitions, one of which (demonstration) is non-linguistic. Types 1 and 3 ("in terms of another science" and "circumlocution") are in reality the same. Type 4 (translation) is more or less followed by F.W. Householder, Jr. in his "On Linguistic Primes" Word 15 (1959) 231f.

19. The synonymies prepared for an assortment of languages, while they may suffer from a variety of malpractices, can be of some use.

20. Others in daily use by myself are ice-cold, red-hot, snow-white, pitch-black, bone-dry, brand-new, dog-tired. A related type (not substituting for very) is found in sky-blue, knee-deep, shoulder-high (or waist-), etc.

21. My wife reminds me of a locus classicus in (pseudo-)Demetrius' On Style (ed. W. Rhys Roberts, 1902) where (I.79, p. 109) we read "Not all metaphors can ..be used convertibly... Homer could call the lower slope of Ida its 'foot', but he could never have called a man's foot his 'slope'".

22. Such indices a tergo are now available for Latin, ancient Greek, Biblical Hebrew, and Russian. For other languages the best that compilers have been able to produce are the deficient "rhyming" dictionaries. Rhyming indices for specific verse texts (based on actual rhymes and not on eye-rhymes or misleading spellings) are better, but they do not represent colloquial or prosaic language.

23. This fact should not worry anybody, since, if a lexicographer is justified in atomizing normal speech-flow into the convenient units of words, he may do the same with the cultural complex underlying the lexical symbols. Furthermore, he may use the same judgment in inclusion or exclusion of cultural items as he does for word-entries.

24. The Linguistic Atlas worksheets are geared primarily to the rural vocabulary. Since, in this country at least, the urbanites and suburbanites have become a sizeable percentage of the population, non-rural variants can be given more attention in our check-list, for instance ordinary but easily overlooked equivalents like rubber-band : rubber-binder, stuffing (in a turkey) : dressing, dish-rag : dish-cloth, etc. But some judgment must be exercised to avoid annoying batches like couch, sofa, davenport, divan, lounge, settee, chesterfield.

25. Some major semantic regroupings are listed here in no particular order. There should be two separate lists for flora/fauna, namely an Arctic-Temperate list, and a Tropical/sub-tropical list; food& drink, food-preparations, tools & weapons, natural phenomena, topography & geography, body-parts (and those of animals), kinship, social status and occupation, colors, size shape direction and position, clothes & ornament, religion, metals and natural substances, excretions, time (& its parts), amusements (sports, games), buildings, the house and its parts, furnishings and furniture, folklore and folkliterature, manufactures & products, political and administrative terms, business & trade, ailments, cures and physical conditions, military & naval, fishing & the sea, farm and farming.